引用格式：王平辉, 裴红斌, 赵俊舟, 等. 网络社会现代治理的挑战与对策. 中国科学院院刊, 2022, 37(12): 1686-1694.
Wang P H, Pei H B, Zhao J Z, et al. Challenges and measurements for governance of modern cyber space society. Bulletin of Chinese Academy of Sciences, 2022, 37(12): 1686-1694. (in Chinese)

网络社会现代治理的挑战与对策

王平辉[1] 裴红斌[1] 赵俊舟[1] 秦 涛[1] 沈 超[1] 刘东亮[2] 管晓宏[1*]
1 西安交通大学 智能网络与网络安全教育部重点实验室 西安 710049
2 西安交通大学 法学院 西安 710049

摘要 信息技术的迅猛发展形成了空前繁荣的网络社会，释放了社会交互中蕴含的巨大生产力。同时，网络社会中出现的电信诈骗、隐私泄漏、网络公害、算法歧视等问题，对社会秩序与安全带来了新的挑战。为寻找网络社会治理途径，推动国家治理现代化建设，文章首先在用户身份治理、网络行为治理和算法治理3个典型场景下分析网络社会中遇到的新问题，及其对经济、社会与安全造成的风险，并对问题治理中所遇到的关键挑战进行剖析，进而从芯片、系统和算法3个层面提出网络社会现代治理的技术对策。

关键词 网络社会治理，身份治理，算法治理，网络安全芯片

DOI 10.16418/j.issn.1000-3045.20221117002

大数据与信息技术的迅猛发展及其与物理世界的深度融合，形成了空前繁荣的网络社会。网络社会中个体紧密联系，交互频繁，呈现出开放、高效、交互、个性化等特征。网络社会重塑了传统社会中的通信与交互方式，释放了社会交互中蕴含的巨大生产力。与此同时，网络社会中出现了电信诈骗、隐私泄漏、网络暴力、网络公害、算法歧视等种种问题，给社会秩序与安全带来新的挑战。为探索如何应对该挑战，寻找网络社会治理途径，推动国家治理现代化建设，本文在3个典型的网络社会场景下，分析与阐明网络社会治理关键挑战，并从芯片、系统和算法3个层面提出网络社会现代治理的技术对策。

1 网络社会治理的挑战

本文在用户身份治理、网络行为治理和算法治理3个典型场景下，首先分析网络社会中遇到的新问题，及其对经济、社会与安全造成的风险，并进一步阐明治理这些问题所遇到的关键挑战。

1.1 用户身份治理面临严峻挑战

*网络用户信息泄露事件层出不穷。*在享受互联网带来的便利的同时，用户也将个人信息上传到了互联网中。这些信息往往带有使用者的隐私，一旦出现数

*通信作者
资助项目：中国科学院学部院士咨询评议重大项目（2022-ZW14-Z-027）
修改稿收到日期：2022年11月12日

据泄露，会产生严重影响。隐私数据必须满足“有限公开”原则，即只有授权用户——通过身份认证的用户，才能搜索到授权允许访问的信息[1]。例如，2014年，国内两个大型物流公司的内部网络系统遭到黑客攻击，1 400多万条快递信息遭到泄露；迄今为止最大的数据泄露事故发生在2013年8月，黑客获取了30亿雅虎用户的姓名、出生日期等隐私信息，给用户带来了极大的安全隐患。攻击者可以利用被泄露的隐私数据，使用被攻击者的身份在网络上发布传播消息，以获取巨大的利益，造成严重的影响。

*现有网络访问控制身份认证技术存在重大安全隐患。*目前最常见的身份认证技术为静态口令认证技术，每一个用户都拥有一个用户名/密码。当用户申请访问时，系统对用户的用户名和对应的密码进行验证，这种方法简单且有效[2]。但随着互联网的发展，这种简单的方法也开始显现出弊端。算力的提升使得攻击者可以对密码进行穷举攻击，这一攻击方式在用户个人隐私信息泄露时会变得更加有效——攻击者可以通过用户的隐私信息推测用户的密码，从而降低攻击者的时间成本。同时有调查显示对于大量不同的账户，绝大多数用户选择使用相同或者数量远小于账户数量的静态密码[3]，因此如果攻击者成功攻击了某用户的账户，那么该用户的其他账户大概率受也会到攻击者的攻击。动态口令技术是由静态口令技术演变而来的，相较于静态口令其具有更好的安全性，能够短暂地维护用户的信息安全。除此之外，若软件系统本身存在漏洞，攻击者也可以通过该漏洞获取访问存放验证口令的服务器的权限或能够直接访问某用户，如黑客对智能汽车、智能家居设备的入侵事件也屡屡发生。因此软件本身可能存在的漏洞也是在设计身份认证技术中所需要考虑的问题。

*用户身份虚拟伪装技术演进迅猛。*在没有强有力的用户身份认证与管理手段的条件下，当用户的信息出现泄露时，攻击者可以利用泄露的隐私信息，访问被攻击者的账户以达成攻击目的。相反，当用户数据没有泄露时，攻击者也可以通过模拟伪造用户的数据来实现攻击。现有的研究工作[4,5]表明，攻击者可以通过对抗样本生成伪装的用户声音，通过伪基站来伪造虚假的通信身份信息及通过对抗攻击的方法生成能够误导身份识别系统的图像等。冒充公检法来进行电信诈骗就是一个典型的例子，攻击者利用特殊计算机软件，通过伪造、冒充执法人员身份的方式，向被攻击者索要钱财以达到自己的目的。在2017年中央电视台3·15晚会中，主持人现场演示了通过一张生成的照片并通过了人脸识别系统。除此之外，攻击者通过特殊的网络软件伪造声音，伪造特殊身份发送短信、拨打电话、生成图片的事例屡见不鲜，因此加强网络用户身份认证及规范行为势在必行。

综上，无论用户数据是否存在泄露的情况，攻击者总能采取手段伪造用户身份，以谋取自身的利益。因此我们强调网络用户身份治理的重要性，将其看作网络社会发展中一个重要且亟待解决的问题。上述问题主要存在于软件层面，而硬件方面存在的问题也会影响用户数据的隐私，进而产生一系列的信息安全问题。

1.2 网络行为治理面临紧迫挑战

网络空间在给人们带来自由表达的便利的同时，网络空间中的个体及群体行为也正变得异常庞杂，而且泥沙俱下、良莠不齐。不良网络内容泛滥、网络违法犯罪活动猖獗、网络空间对抗与博弈等违法违规网络行为已经愈演愈烈，严重威胁网络空间与现实社会的安全稳定与健康发展。

*不良网络内容泛滥。*互联网与数字技术的蓬勃发展，一方面带来了丰富的信息内容，另一方面也带来了虚假谣言、低俗拜金、暴力反动等不良网络内容的泛滥。虚假谣言信息内容在网络空间中大肆传播、滋长蔓延会严重影响社会秩序，已经成为一大公害。短视频平台和直播平台存在大量低俗内容，利用主播

搭讪、诱惑打赏等低俗手段博人眼球，传播如种族歧视、性别歧视、拜金主义等不良内容。网络游戏平台存在大量血腥、暴力、恐怖等不良内容，宣扬暴力极端主义、恐怖主义等极端思想。这些不良网络内容利用钻法律空子、打擦边球等方式诱导大众点击与关注并从中牟利，对网民尤其是未成年人和青少年人形成不良导向，影响身心健康。

*网络违法犯罪活动猖獗。*网络空间中的网络诈骗、网络赌博、网络传销、网络非法集资等违法犯罪活动十分猖獗。中国互联网络信息中心统计发现[6]，截至 2022 年 6 月，我国 17.8% 的网民曾经遭遇过网络诈骗，较 2021 年增加了 1.2 个百分点，其中占比最多的网络诈骗包括虚假中奖信息诈骗、网络购物诈骗、冒充好友诈骗和钓鱼网站诈骗等。中国司法大数据研究院分析了从 2017 年 1 月至 2021 年 12 月我国信息网络犯罪案件的特点和趋势[7]，发现网络诈骗罪案件量占比最高，达到 36.5%。这些网络诈骗案件多以办理贷款、冒充他人身份、虚假招聘等方式或话术来欺骗受害人，给受害人财产造成重大损失，而且目前网络犯罪案件量呈现出逐年上升的趋势，需要格外重视。

*网络空间虚假信息传播与博弈。*网络空间已经成为领土、领海、领空之外的第四种主权疆域。网络“水军”、网络公关等产生的虚假信息传播与博弈已经成为常态。网络平台的虚假实体、虚拟用户背后的操控系统能够堂而皇之利用各种技术手段，系统传播虚假信息、谣言、偏见，实现舆情对抗与博弈，已经成为危害网络空间内容安全的重大威胁。

1.3 算法治理面临多维度挑战

近年来，以机器学习算法为主导的人工智能技术已成为各大网络平台的核心驱动力，其在加速信息传播、便捷百姓生活、繁荣数字经济和促进社会发展等方面发挥了重要作用。与此同时，由于在公平性、透明性和安全性方面的缺陷，机器学习算法的不合理应用对网络社会的正常秩序带来了多维度挑战，影响到社会公平公正、网民合法权益及虚拟财产安全等诸多方面。算法的应用风险和治理挑战已成为各界关注的焦点。

*算法歧视普遍存在。*基于大数据与机器学习算法的自动决策系统被广泛用于提供个性化的辅助决策，例如电影推荐、贷款申请、广告投放、求职和约会等。算法歧视是指由数据和算法的偏向性导致的决策偏见。一种典型的算法歧视被称作大数据“杀熟”，其表现为同物不同价，平台算法依据大数据推断用户的价格敏感度，并针对性地进行差异化商品标价。北京市消费者协会 2022 年发布的《大数据“杀熟”问题调查报告》显示超六成受访者曾遭遇过大数据“杀熟”，而购物、旅游与外卖类网络平台是大数据“杀熟”的重灾区。网络平台算法也常存在性别与种族偏见。2019 年，脸书公司开展了一次算法中立性测试，测试通过算法向全球投放科研岗位的招聘广告，在全部人为环节都保证性别中立的前提下，算法最终产生了有性别偏见的广告展示结果[8]。可以看出，网络平台中的算法歧视不仅损害用户的合法权益，同时影响社会应有的公平公正，扰乱社会正常秩序。造成算法歧视普遍存在的主要原因是算法的价值预设和数据的采集偏向：首先，算法由网络平台所设计和研发，本身蕴含预设的价值偏见，并非完全客观；其次，算法依赖的训练数据往往并非中立，数据采集范围存在偏向，造成算法异化，形成算法歧视。

*算法安全风险日益凸显。*研究表明深度神经网络等算法存在内生的脆弱性，攻击者能够利用微小的对抗扰动或后门攻击，在物理世界实现对算法决策的任意操纵[9]。对抗鲁棒性和泛化能力等表达能力的缺陷，导致了深度学习算法在应用领域中，尤其是自动驾驶、智能医疗等容错率低的关键应用中的潜在安全隐患。例如，人脸识别算法具有高准确率、高效率、非接触性识别等优点，成为网络平台身份验证的主流方式之一。2021 年 RealAI 公司利用基于算法漏洞特制

的眼镜，成功欺骗了19款手机搭载的人脸识别算法。另一项研究显示，将一张用简单打印出的涂鸦贴画贴在真实的路牌上，就可以“迷惑”人工智能自动驾驶系统，使其做出错误决策[10]。随着深度学习技术的广泛应用，算法安全隐患带来的风险日益凸显，逐渐成为制约网络平台与人工智能技术发展的瓶颈。

*“黑盒”算法监管困难重重。*平台算法治理的另一挑战来自监管困难。一方面，网络平台用户的所见只有算法输出，难以监督算法的决策过程。另一方面，由于大多算法属于不透明的“黑盒”算法，人类难以理解算法究竟如何做出决策，平台监管者同样难以判断算法决策是否公平、合理、可信赖。部分网络平台还可能以商业秘密为由，逃避外界对算法的监管。打开“黑箱”算法，将算法隐患与风险置于在监管视野之内，是算法治理的基础和必要前提。

2 网络社会治理的技术对策

2.1 基于安全芯片与管控系统的治理对策

由于移动办公已经成为日常工作模式，政府、医疗、金融等并非最高安全等级的部门与行业，必须要在外部访问单位的业务内网处理日常业务，需求量巨大。鉴于内网高安全性要求，常用的安全网关、虚拟专用网络（VPN）等网络安全技术无法保证内外网访问的安全。

*要从根本上快速解决和缓解上述网络安全问题，必须从网络安全芯片底层设计做起。*一方面保护数据资源免于非授权访问、篡改，另一方面，建立芯片级安全通信专用通道，实现数据和通道两个方面的芯片级策略化安全隔离。网络安全芯片底层设计的主要优势在于增加系统安全性的同时，极大地提高数据处理效率，策略化的安全隔离，规避了以易用性为代价的简单物理隔离。具体而言，使用纯软件的方法实现网络安全加解密、身份认证、网络入侵检测等算法存在较多缺陷，如执行各类算法的时间较长、资源消耗较大，且难以实现密钥等机密资源的安全存储。因此，集成一种或多种密码算法的集成电路芯片，即安全芯片，可从硬件层面来解决上述问题，在增加系统安全性的同时，可极大地提高效率。

传统的网络安全芯片防护如图1所示，这种工作模式虽然能够实现对网络传输的加密，在硬件层面保护敏感数据通信安全，但系统本身容易受到网络攻击，很难保障自身安全。针对以上问题，我们提出了新型网络安全管控芯片和系统的架构[11]，主要包含数据安全单元、网络安全管控单元与网络数据处理单元，如图2所示，实现敏感数据的隔离、网络访问控制、安全模式热切换、身份的认证校验。网络数据处理单元，与加解密算法控制引擎进行通信，利用加解密算法控制引擎对当前数据包中的载荷进行加解密，并完成网络协议中的校验。

如图3所示，基于网络安全管控芯片，研发基于安全管控芯片的安全智能终端（例如移动终端），构建内外网安全管控的软硬件生态，实现内、外网安全的全方

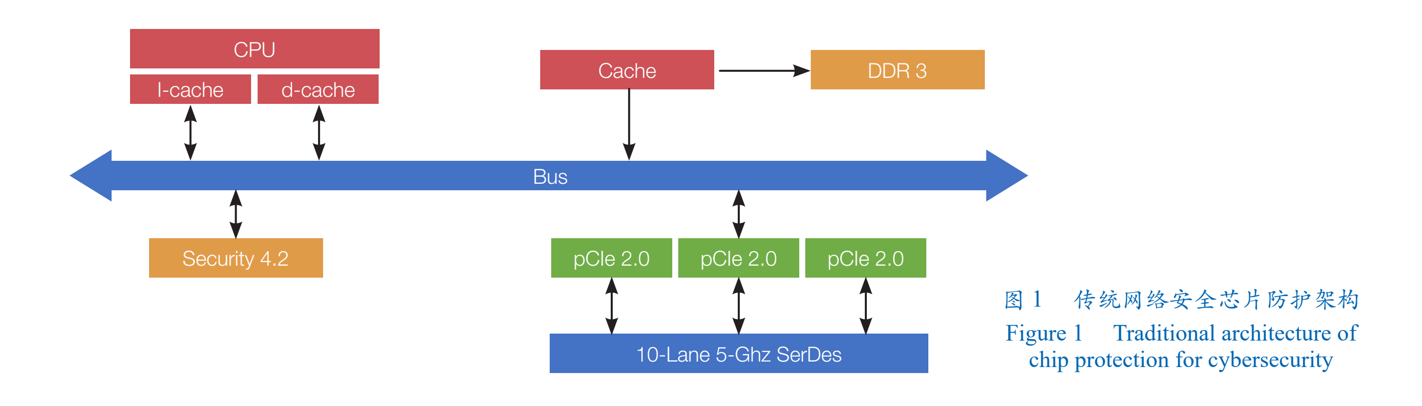


图1 传统网络安全芯片防护架构
Figure 1 Traditional architecture of chip protection for cybersecurity

位管控，进而形成完全自主可控的产业生态。通过以上软硬件技术，可以保证使用新型网络安全芯片的智能终端既可以自由访问外网，也可以安全访问内网数据，同时确保内网数据不会通过设备泄露到外网。

2.2 系统与应用层治理对策

网络身份和网络行为治理困难的主要难点在于网络空间中与身份和行为相关的数据类型多样化、数据非结构化、数据规模庞大、数据动态性、数据呈孤岛状分布且不流通等，导致难以对网络实体身份和网络行为进行表征、识别与关联，缺乏有效的检测技术，从而难以识别用户身份并阻断违法违规网络行为。近几年来随着大数据分析、自然语言理解、数据联邦、

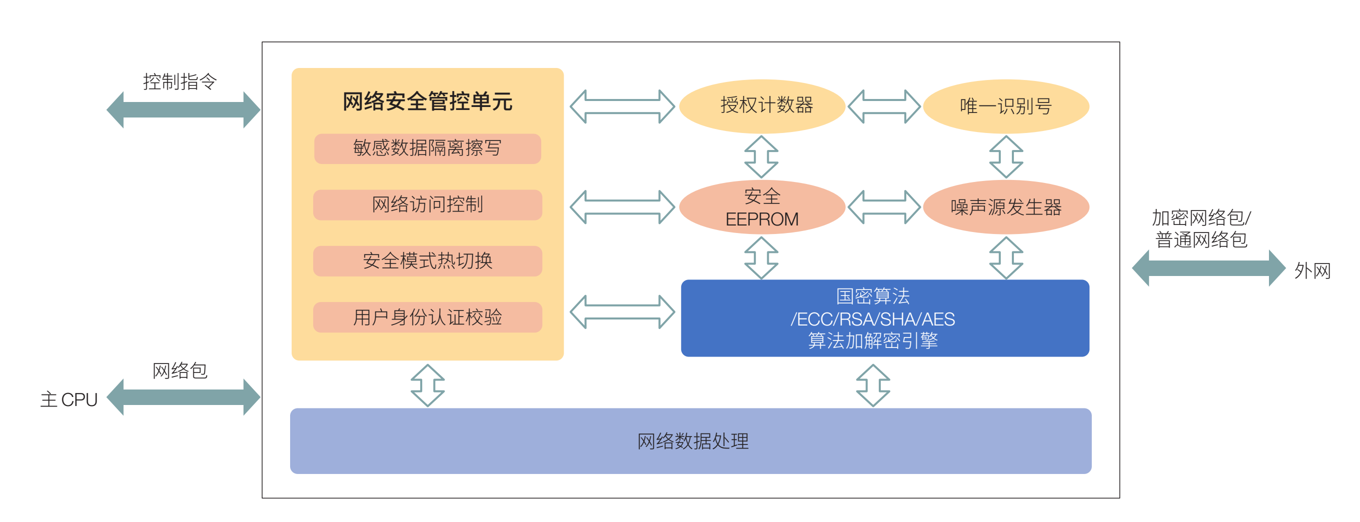


图 2　新型网络安全管控芯片架构

Figure 2　Novel architecture of chip control for cybersecurity

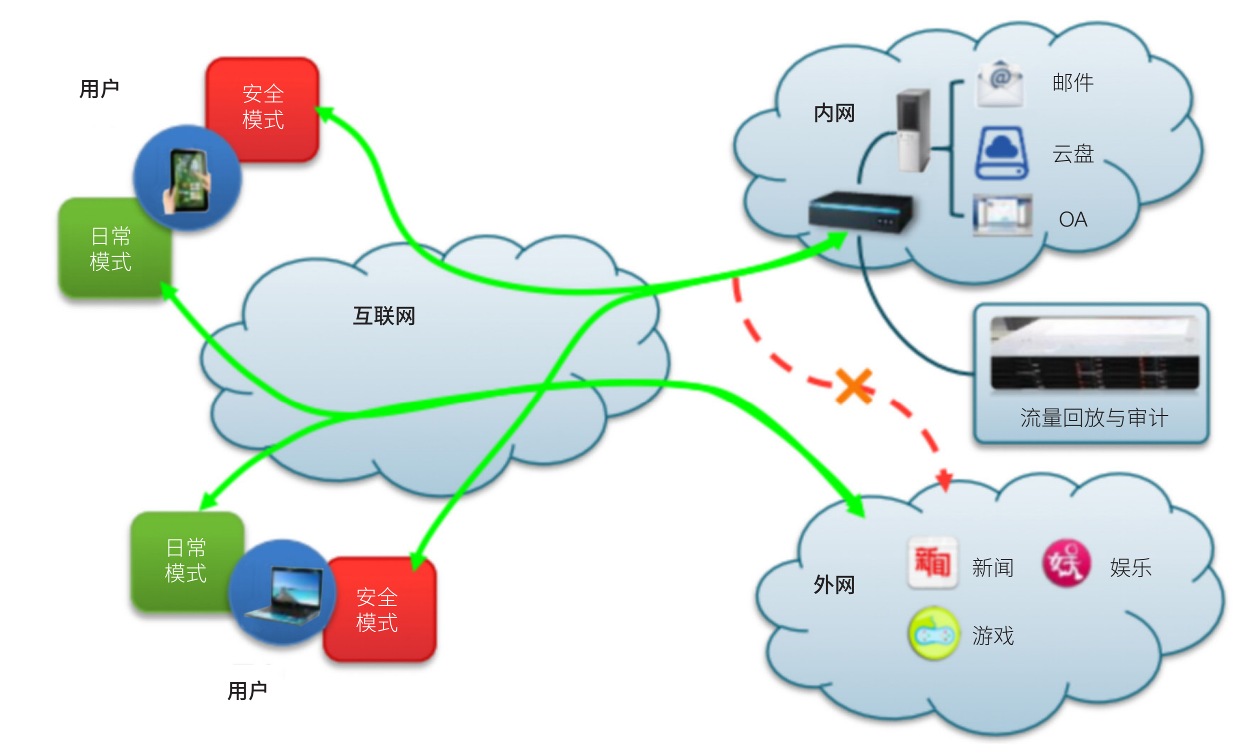


图 3　基于网络安全管控芯片的内外网安全管控系统软硬件生态

Figure 3　Software and hardware ecology of internal and external network security management and control system based on cybersecurity management and control chip

多方计算等信息技术的快速发展，为网络身份和网络行为治理思路带来新的启示。

通过制定网络身份表征标准并建立多方联合计算与数据流通框架解决网络身份识别、关联与认证问题。构建针对网络实体的标准化身份表征类似于给每个网络实体发放一个身份证，该身份证记录了网络实体的各种属性，并且得到权威部门的可信认证。标准化身份表征可以有效解决网络身份识别、关联与认证问题。构建标准化身份表征需要借助多模态机器学习技术融合多种类型数据进行属性推理[12]，利用小样本学习与迁移学习技术解决标注样本稀少问题[13]，利用数据联邦与多方计算技术解决数据孤岛与不流通问题[14]等。

通过制定网络行为表征标准并建立多方联合计算与数据流通框架解决网络行为识别、检测与治理问题。网络行为治理可以借鉴网络身份治理的思想，构建标准化网络行为表征，解决网络行为治理问题。在进行网络行为表征时，由于网络行为具有较强的动态性和阶段性，处于不断发展变化中，因此需要额外考虑时效性。为解决该问题，可以利用数据流实时分析技术及时决策和推理[15]，也可以通过构建事件知识图谱以厘清网络行为的阶段性变化特征。

2.3 算法层治理对策

发展算法的可解释性是算法层治理的核心。可解释性是打开“黑箱”算法的钥匙，算法决策机理的透明化是监管与治理算法的先决基础，是避免算法歧视与风险，保证决策可信、正当和理性的重要约束。算法的解释方法主要包括两大类：① 事前解释，一般指具有高度透明性的自解释算法，可直接被用户查看和理解，如常用的决策树、逻辑回归算法等。在高风险的领域中，如医疗辅助系统，应当使用这类自解释算法。② 事后解释，一般指使用解释性工具对半透明、不透明的复杂算法进行解释。例如，归因工具是一类常用的解释性工具，它的解释可以回答算法决策是依据哪些目标特征做出的，并可精细地量化每个目标特征对最终决策所作出的贡献度；另一类博弈交互工具的解释可以回答算法建模了哪些类型的知识点，进而可根据这些知识点的质量，解释算法的各种安全性指标，如在对抗攻击下的算法鲁棒性，或在新的测试环境下的算法稳定性等。

发展第三方算法审计机构是算法治理的有效途径。其目的是对平台算法是否合法、合规进行客观鉴定。算法审计机构可以由政府监管部门牵头，协同行业自律组织、专家学者等组成，也可按照市场化原则培育独立的第三方算法审计机构从事该项工作。具体算法审计流程为：① 第三方组织根据委托审计主体的要求，制定专门的审计目标与实施方案。② 审计人员对平台的系统配置和算法设计文档、运行记录等进行查验，通过人工或自动测试工具进行技术测试，获取相关信息，进行分析取证。③ 审计人员以法律法规、行政部门指导制定的程序合规标准体系为依据对审计证据进行审计评估，并根据国家强制性标准、推荐性标准区分不同的合规程度。④ 审计人员提出初步审计意见与程序合规建议，并与被审计平台就其合理性与必要性进行沟通交流，审慎得出审计结果并出具审计报告，作为平台合规性验证和归责处罚的依据。算法审计应把握以下原则：① 坚持持续、动态审计。随着平台算法的更新演化，可能出现难以预知的增量风险，因此应持续开展动态算法审计。② 遵循分类、分级原则。应根据平台算法性质及运行环境制定相应的审计方案，对重点领域的高风险算法要加强审计。③ 恪守保密原则。审计机构和审计人员应当对涉及平台算法模型的训练数据集、用户数据等构成平台核心竞争力的商业秘密严格予以保密。

3 网络社会现代治理的政策建议

随着信息化水平逐步提升，信息技术嵌入网络社会治理的身份治理、行为治理和算法治理等各层面各

环节，与法律、市场、平台自治规则一起发挥着重要作用。现代网络社会治理仍存在严峻挑战，需要多方协同努力，在此提出3点建议。

（1）加快研发网络安全管控芯片和系统，突破我国网络社会治理中的“卡脖子”问题。网络安全技术前期焦点主要集中在软件和应用开发，而网络安全硬件平台作为其载体，国产化产品的大规模应用即将进入加速期，国家和市场关注度及重视程度正在快速提升。党的二十大报告中提出“加强企业主导的产学研深度融合，强化目标导向，提高科技成果转化和产业化水平。强化企业科技创新主体地位，发挥科技型骨干企业引领支撑作用，营造有利于科技型中小微企业成长的良好环境，推动创新链产业链资金链人才链深度融合”，建议我国企业发挥市场主体的优势作用，积极联合高校、研究院所等开展网络安全芯片及硬件平台相关理论和技术的研究，强化在网络安全芯片与管控系统、安全可信操作系统、密码算法与硬件平台技术等方面的国产化管理，实现全链条的自主可控。明确网络安全芯片关键产品国产化、密码算法国产化与网络安全硬件平台国产化的政策要求，制定关键产品审查备案制度，组织产业链调研，适时推出国产化扶持政策。

（2）面向网络/电信诈骗等重大国计民生问题，发展隐私计算技术，建设安全的数据共享平台，推进平台间的数据共享。网络/电信诈骗等网络社会不良行为常采用技术手段隐匿行迹，难以在单一数据源中被及时地检测和发现。我们建议在政府监管下，建设安全的数据共享平台，在保护用户隐私和防止数据滥用的前提下，整合多方数据源，实现对隐蔽不良行为的有效探测和发现。建立安全的数据共享平台应当满足以下要求：① 制定统一的数据标准和接口，明确多方数据共享模式，支撑数据的高效共享与交换；② 建立数据分级管理体系，合理确定共享范围，防止数据滥用；③ 发展隐私计算技术，防范隐私数据泄露，保障共享数据安全；④ 确定共享数据的权利与义务，保障数据共享体系的长效运行。

（3）建立第三方算法审计制度，培育第三方算法审计专业机构。鉴于算法审计技术的复杂性，在政府监管过程中，可以由监管部门牵头，协同行业自律组织、专家学者等组成第三方，对网络平台算法设计的公平性、可解释性、安全性等开展审计。长远来看，应当按照市场化原则培育独立的第三方算法审计机构（类似专业的财务审计机构）承担该项工作。对此，可以采取分步走的办法，先由相关部门如工业和信息化部、中央网信办、国家市场监督管理总局、司法部等部门会商研拟方案，鼓励若干信誉良好、有实力的司法鉴定中心在内部设立算法审计部门，接受政府或司法机关的委托，从事算法审计和鉴定工作。待市场条件成熟、有足够的市场业务容量后，鼓励成立专门的算法审计机构，作为独立的市场运营主体，开展算法审计业务。即通过算法审计，实现“借力打力”，由技术引发的问题还通过技术手段来解决，促进网络社会的“科技向善”。

参考文献

1 房梁，殷丽华，郭云川，等. 基于属性的访问控制关键技术研究综述. 计算机学报, 2017, 40(7): 1680-1698.
Fang L, Yin L H, Guo Y C, et al. A survey of key technologies in attribute-based access control scheme. Chinese Journal of Computers, 2017, 40(7): 1680-1698. (in Chinese)

2 冯登国，陈伟东. 基于口令的安全协议的模块化设计与分析. 中国科学（E辑：信息科学）, 2007, 37(2): 223-237.
Feng D G, Chen W D. Modular design and analysis of secure protocol based on passwords. Science in China (Series E: Information Sciences), 2007, 37(2): 223-237. (in Chinese)

3 Florencio D, Herley C. A large-scale study of web password habits// WWW '07: Proceedings of the 16th International Conference on World Wide Web. New York: Association for Computing Machinery, 2007: 657-666.

4 Tang S Y, Shu X M, Shen S F, et al. Study of personnel

positioning in large area based on pseudo base station. Procedia Engineering, 2014, 71: 481-485.

5 Chen M Y, Lu J D, Wang Y, et al. DAIR: A query-efficient decision-based attack on image retrieval systems// SIGIR '21: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: Association for Computing Machinery, 2021: 1064-1073.

6 中国互联网络信息中心. 第50次中国互联网络发展状况统计报告. (2022-08-31). http://cnnic.cn/n4/2022/0916/c38-10594.html.

China Internet Network Information Center. The 50th statistical report on China's internet development. (2022-08-31). http://cnnic.cn/n4/2022/0916/c38-10594.html. (in Chinese)

7 中国司法大数据服务网. 司法大数据专题报告之涉信息网络犯罪特点和趋势（2017.1-2021.12）. (2022-08-01). http://data.court.gov.cn/pages/reportshow.html?filename=司法大数据专题报告之涉信息网络犯罪特点和趋势.pdf.

China Justice Big Data Service Platform. Judicial big data on features and trends of Internet criminals (2017.1–2021.12). (2022-08-01). http://data.court.gov.cn/pages/reportshow.html?filename=司法大数据专题报告之涉信息网络犯罪特点和趋势.pdf. (in Chinese)

8 Lambrecht A, Tucker C. Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. Management Science, 2019, 65(7): 2966-2981.

9 Kurakin A, Goodfellow I J, Bengio S. Adversarial examples in the physical world// Kurakin A, Goodfellow I J, Bengio S. Artificial Intelligence Safety and Security, 2018: 99-112.

10 Liu A S, Liu X L, Fan J X, et al. Perceptual-sensitive GAN for generating adversarial patches// Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2019: 1028-1035.

11 王平辉, 杨晨, 管晓宏. 一种用于网络安全芯片的内外网访问控制方法与该网络安全芯片: 中国, CN202011616459. 2022-03-20.

Wang P H, Yang C, Guan X H. Network secure chip based inner-outer access control method and network secure chip: China, CN202011616459. 2022-03-30. (in Chinese)

12 Tsimpoukelli M, Menick J L, Cabi S, et al. Multimodal few-shot learning with frozen language models// Proceedings of the Advances in Neural Information Processing Systems. New York: Curran Associates, Inc., 2021: 200-212.

13 Lan L, Wang P, Du X, et al. Node classification on graphs with few-shot novel labels via meta transformed network embedding// Proceedings of the Advances in Neural Information Processing Systems. New York: Curran Associates, Inc., 2020:16520-16531.

14 Kairouz P, McMahan H B, Avent B, et al. Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 2021, 14(1-2): 1-210.

15 Hartvigsen T, Sen C S, Kong X N, et al. Recurrent halting chain for early multi-label classification// KDD '20: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: Machinery, 2020: 1382-1392.

Challenges and Measurements for Governance of Modern Cyber Space Society

WANG Pinghui[1] PEI Hongbin[1] ZHAO Junzhou[1] QIN Tao[1] SHEN Chao[1] LIU Dongliang[2] GUAN Xiaohong[1*]

（1 MOE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China;

2 School of Law, Xi'an Jiaotong University, Xi'an 710049, China）

Abstract The rapid development of information technology has unprecedentedly created a prosperous cyber society and greatly enhanced productivity facilitated by social interaction. At the same time, many problems emerge in the cyber society, such as telecom fraud, privacy leakage, Internet pollution, and algorithmic discrimination. The problems bring new challenges to social order and security. In order to find the way of cyber society governance and promote the modernization of national governance, this paper first presents the analyses on the new problems encountered in the cyber society in three typical scenarios, i.e., identity governance, behavior governance, and algorithm governance, as well as their risk to the economy, society and security. Furthermore, this paper lists the key challenges for the problem governance, and presents technical countermeasures for modern governance of cyber society from three levels of chip, system, and algorithm.

Keywords cyber space governance, identity management, algorithm management, cyber security chip



王平辉 西安交通大学网络空间安全学院副院长、教授、博士生导师。主要研究领域包括大数据与网络安全。E-mail: phwang@mail.xjtu.edu.cn

WANG Pinghui Deputy Dean and Professor of the School of Cyber Science and Engineering, Xi'an Jiaotong University. His main research areas include big data and network security. E-mail: phwang@mail.xjtu.edu.cn



管晓宏 中国科学院院士，西安交通大学电子与信息学部主任、教授、博士生导师。主要研究领域包括网络化系统的优化与安全。E-mail: xhguan@xjtu.edu.cn

GUAN Xiaohong Academician of the Chinese Academy of Sciences, Director and Professor of the Faculty of Electronic and Information Engineering, Xi'an Jiaotong University. His main research areas include the optimization and security of networked systems. E-mail: xhguan@xjtu.edu.cn

■责任编辑：张帆

*Corresponding author